Slide 2:

To date based on the pilot findings, tell you a little bit about the analysis that we did over the last year or so, partly from the pilots and then also through another piece of analysis that NREL did for us using model data and also utility bill data. And then, finally I will tell you a little bit about our next steps and what were are planning as we move forward with implementation, both in terms of implementation, but with an emphasis of what we plan to look at further as we proceed.

Slide 3-4:

OK, so some of you may have already heard of these or seen these many times, these are the guiding principles DOE came up with when we put together the Home Energy Score program and we look back at these as kind of our guiding posts to make sure we are still in line with what it is we were trying to achieve initially when we created the program, and I am happy to say that I think for the most part we have been able to fulfill these principles and adhere to them pretty well. The last one about quality assurance, we haven't tested that as closely and it is something that we are going to look at further in this first phase of implementation. We just didn't have the time and range of scoring that we were able to do, in terms of that in the first phase, during the pilot phase. That is something we will look at further.

Slide 5:

So again, giving you a really quick background. The basis for the Home Energy Score is done by collecting about 40 data points in the home and looking at the envelope and the major energy systems, being heating, cooling and hot water. We make some assumptions about how the home is being operated in terms of people that live there, their behavior, and how they set their thermostat, etc. The size of the home does matter, it is used to, the volume is basically used to calculate what the load requirements will be for that home, in addition to obviously looking at where it is located in the country. We do not adjust the 10-point scale for size. The 10-point scale is, however, adjusted for the local climate and we went from 19 climate zones that we used to use for the pilot phase to adjusting that scale for 248 different weather stations around the country, at this point.

Slide 6:

Before we announced the Home Energy Score we did a fair amount of analysis and focus groups, etc., to create the Home Energy Score and the Scoring Tool and then we've done analysis since that time during the pilots and beyond. So, I am going to go through quickly what we found in all that analysis and stress the major findings.

Slide 7:

Before the pilots, that is in 2010, we did hundreds, I guess thousands I should say, of simulations where we looked at different variables, in terms of asset variables that can affect how much energy we are using in the home. So, not looking at how you set your thermostat, or how many loads of laundry you do, but looking, or how many TVs you have, but looking at the assets—the envelope, the windows, the roof, the foundation, insulation levels, the ceiling, whether the ducts are sealed. Things like that, in addition to heating, cooling and hot water, and we isolated those factors that are asset factors and we believe, those we found contribute 2% or more to overall energy use. I should say there are certain things we excluded, like pools, there were just various reasons we decided to do that but that is something we might rethink at a future date. But obviously, pools and hot tubs do contribute, often more than 2%, but it also depends a lot on the behavior of the occupants and how they are using those assets, so that is why we, at that point, left those out, or have to date.

We also found through our focus groups and also through a lot of social science research that had already been done by others, that people are often or frequently motivated by their peers and not necessarily only motivated by dollars, or comfort, or other things they might think about as motivational, that they actually often affected in a kind of illusory way by just what is going on around them. We also found in our focus groups that people responded much better to information that was clearly presented, which seems obvious, but also that was just not an overload of information, that they tended to get distracted if you have them too much to look at upfront. That's not to say you can't give them additional information at the follow-up, but in terms of grabbing people's attention it made a lot of sense to give them simple graphics and information that they could quickly assimilate.

Slide 8:

In terms of what we studied since starting the pilots in 2011, we were looking again at these guiding principles, at what we could try to prove out, that we are actually achieving these, and seeing how effectively we are achieving these, and we found that when we rescored homes by different assessors, this is not a huge sample but when rescored homes by different assessors, generally they came within 10 percent of each other and only in one case were they more than 1 point off from each other. That is, like if the first person got a 6, the second person got a 5, or a 6, or a 7. So, in general, we thought all of the assessors we were using during the pilots were consistent with each other, in terms of the rescoring number, but that's something we plan to look at further in the next phase. We also found through a modeling effort that, we felt that give the type of scale we had selected, where we don't have a very granular scale, it's only 10 points, we found that if you vary the different factors that are considered asset factors in your modeling that the estimates, even if you are not certain about certain characteristics like wall insulation or attic insulation or how tight the house is, that as long as you were within a reasonable understanding of what's going on in that house, that you could, that you would

score the home within one point of what it actually should be scored and I'll go into what I mean by that a little later.

We also did some analysis, NREL did some analysis where we looked at that the scoring tool predicted, in terms of energy use given certain home characteristics versus what was actually going on in that home, where we did have utility bills and I'll show you how we compared to other tools in estimating energy use. But I should stress that this tool is not intended to estimate energy use, and neither were the other tools we looked at. These were asset tools. With that said, it is sometimes comforting to know you're going to be within a reason range of what is really going on there in reality, that is what people are actually using.

Slide 9:

We found that the Score was easy to understand. We did some additional analysis with homeowners and it was easy to understand. We also found that it can be done in what we think is a reasonable amount of time. Fifteen minutes if you are doing it a part of another effort. An hour or less if you are doing it as a stand-alone effort and we also found through sensitivity analysis that we did not believe we needed to do a blower door test. Again, I'll go into more detail on these in other slides.

Slide 10-11:

OK, so in terms of the program improvements to date, as I said yesterday and you probably heard in the past, we had these pilots around the country.

Slide 12:

They led to a lot of improvements as well as other analysis that we were doing with our experts that led us to make adjustments to the tool and other pieces of the program. So, what they told us was that we needed to provide some additional materials for assessors and for homeowners, so we're in the process of developing those or we've filled some of them. We also found that it would be helpful to even go further in simplifying the graphic that we were using during the pilot, so we did make that a bit simpler. We did some analysis that showed that people tend respond to bigger numbers, so we changed the front page of the score from showing just one year of savings to 10 years of savings. The reason we went with 10 years is we thought that was a reasonable amount of time, plus if people wanted to see what they were going to save in a year it was easy to do the math. We are trying to move forward with some type of customization of the information. That is our intent. It has been a little bit tricky just in terms of technically doing it from the computer side, but we're working towards making that more feasible in the

future. We also found that in some cases of the recommendations we were generating during the pilots were off, because they were looking at, for instance they were looking at if you had, you know, a 20-year-old boiler and you were gonna go replace it, we were calculating the savings as, what you would replace it if you bought an ENERGY STAR boiler or furnace, or whatever, today, versus a off-the-shelf non-ENERGY STAR. It wasn't looking back at the base case of the house. So this was messing things up, both in terms of underestimating savings, if you had a crummy piece of equipment, because it was far worse than what you could get, even if you didn't buy ENERGY STAR, or it was overestimating savings because we weren't taking into account that you already had an ENERGY STAR heater or cooling system. So we've fixed that, and we also found that, in some cases there were some pilot partners that really stressed the fact that they weren't, they were a bit perturbed by the fact that if you had some really poor scoring homes, that they could still not necessarily get to a be a 7, 8, or 9, or 10, that they might only get to be a 6, or in some cases possibly even just like a 5, or a 4, and so they felt that it was important to provide material that would allow people to understand the context of the scale, and that it's, well, it allows you to compare homes, it's also useful as a point to understand that if you're a 2 today and you can get to be a 6, that's still, you know, great, in terms of what's cost effective for that home. We don't have enough data at this point to show you that if you're a 1940s bungalow, you know, the best you're gonna be is a 7, or, I making that up, but that is the kind of information that I think people would find useful; we don't have it yet, ‘cause we just haven't scored enough homes. We did a lot of work with the assessors, the 31 assessors who were involved in the pilots, they gave us tremendous amount of feedback on, kind of what they thought of the scoring tool, and the program in general. We added a lot of information within the scoring tool so you could click on question marks for every single field, and get additional guidance that way to guide the assessor as they're going through it in case they forgot what we're actually asking for in a specific field. They also indicated where we might not have enough choices, or too many choices, and so we, we did make some changes to the scoring tool based on that. We've greatly enhanced our training and our testing to cover not only the scoring tool, but also to, and the scoring program, but also to cover some building science issues as well, just to be as a refresher and to ensure that people have been going into homes, you know relatively recently, and know how to make some basic calculations. We've also expanded the testing to include, not just the multiple choice portion which it was before, we expanded that but we've also included a second part of the test which requires the assessor to make some calculations about homes and then enter that and score some fictitious homes in the scoring tool itself.

Slide 13:

We looked at three different software tools. We looked our own Home Energy Scoring tool, we looked at REMRate, and we looked at SIMPLE, a version of SIMPLE, and all of those were intended to provide an asset, well, I mean they can be just used for other things, I don't want to speak for the other tools, but, in, none of them were intended that the way we were using them to predict energy use in terms of actual utility bills, but we did that kind of analysis where we compared those estimates to actual energy use, just to give us a sense of, you know, grounding, of whether or not these things were way off, or whether they were in a reasonable range.

Slide 14:

OK, so, getting back to, OK, so in terms of the scoring tool improvements, I already mentioned some of the ways that we changed how we were calculating savings and generating recommendations, I think we did some improvements to the user interface, I also mentioned some of that. I not gonna go through every single one of these but we will make these slides available if you're interested in, in seeing the specific details, again there were certain things that assessors indicated as being just either inaccurate or possibly not including enough options for them, and so we made those adjustments. We also adjusted the, the plug load -- can now be scaled by floor area, it used to be just a sort of standard number for every house, and now the lighting and small appliances for the home are adjusted in terms of the load calculations to be scaled by floor area. We did change the standard occupancy assumption based on some research that Danny Parker did. We found that, in general there was a sort of a, a linear distribution, a linear relationship between the number of bedrooms and actually how many people are living in a house, it wasn't a direct one-to-one, but I can make that equation available to people if they want, I think it was like .7 plus .3 times bedrooms or something, and so that's what we used. It didn't change too much, in terms of what ended up being the case of how many people we're assuming are living in a house, but the number of occupants is tagged to the number of bedrooms… or tied to. We made some adjustments to our defaults, and some of that was done based on information that we got out of the RECS 2009 data, as well as information that was updated from the Building America house simulation protocols and benchmark definitions. And I also mentioned that we had gone from 19 REC zones that we were using for climate to 248, and that was partly by, because of something that the folks at CEC in California raised to us and it really it turned out to be important to make that change for the whole country, and I can go through that analysis as well.

Slide 15:

{Webinar Host 1: Will you take a couple questions?}

Sure.

{Webinar Host 1: OK, can you address the decision to go from 1 to 10 with 10 being the best score, opposite of the HERS rating, and the chance that this could create confusion with the various rankings?}

I think that's something that people have struggled with, as to you know, whether we're trying to push net zero, or we're trying to show zero energy use; the reason we went with the 10-point scale is that we

though in general that people outside of the energy world understand a 10-point scale, like, you know, going back to grade school and taking 10-point quizzes, or seeing Bo Derek in the movie "10." Ten was associated with good, or great, and 1 was associated with not-so-great. And that's why we went to that. I don't think there's any, you know, perfect solution to what scale you should use, but that's how we came to that.

{Webinar Host 1: OK, and could you clarify your comments in terms of the other tools not being designed to test against utility bills?}

I think I misstated that, if that's the impression I gave that's not what I meant. I meant that, I was saying that we are not expecting the Home Energy Scoring Tool to predict your utility bills, but we did an analysis where we compared the estimated energy use from the Home Energy Scoring Tool to actual energy use from homes. We did the same thing with two other tools. It's not that they were, you know I'm just saying, it's not intended to do that, but it was, we were kind of trying to do a comparison by looking at three different tools and how well their predicted energy use corresponded to utility bills. That's all, it's just, I was trying not to overstate that ours was not intended and others were, it's, I think in all cases, and maybe Dave can answer this, in all cases those were being used as asset tools, and we held certain assumptions constant about how they were being used, because we did not have the information about who was living in those homes. And Dave, if you want to add anything you can.

{Webinar Host 1: OK, well then, could you clarify what is the accuracy level compared to a real energy bill?}

That's a good question. I think accuracy is all within the eye of the beholder, I guess. What we're trying to achieve is accuracy in terms of relative accuracy. We're not saying that your home is going to use this much energy regardless of who lives in it. Obviously, if you're a person who leaves the lights on and, you know, does three loads of laundry every day, and you know, washed the dishes constantly, you're gonna have a higher energy bill than somebody who's more careful, and that's just the way it is. So we're not trying to predict how much energy is going to be used by a specific homeowner, so we're not accurate in that sense. What we are accurate, and I think we can feel very confident about this is that if you're, if you take the same homeowner and you put them in two different homes, if one home has a not-so-great heating system and cooling system and it's not very well insulated, that homeowner is going to use more energy in that home than in the one that scores, if that home scores a 3, and another home scores a 7, if they live in the 7 and they live the same way as if they were going to live in the 3, they will use less energy. So it gives you a point of comparison between homes, but it does not predict the actual utility bills, so there's no accuracy involved in that sense. It's accurate in terms of scoring homes in a relative

sense. A crummy home will get a poor score, a medium home will get a medium score, and a really good home will get a good score.

{Webinar Host 1: OK, thanks, Joan.}

OK, so in terms of going from 19 climate zones to the 248, the reason we did that is that we, we took, I have to remember how we did this, we basically took the same home and used different zip codes within, like let's say we were looking at Zone 1 from the 19 climate zones we used to have from RECS, I believe that was like the NorthEast; so if we took two different zip codes from Maine and Massachusetts, if they were all on the same climate zone, and we modeled the same home in those two different zip codes, what we found was they wouldn't necessarily score the same, and obviously that was a problem, because it turned out that even if you were within the same REC zone, which was, you know, a pretty big zone in some cases, you'd have different weather obviously if you're in Bangor, Maine, then if you're on the Cape, or if you're in western Mass., or, wherever you are; you could be in the same REC zone but you have different weather, and the load calculations done by the tool are based on the local weather, but previously we'd only created 10-point scales that were adjusted to 19 different areas of the country. Now they are adjusted based on the local weather so, I should say it's only based on the 248 weather stations associated with PMY2 data, now we've moved, or the country's moved to PMY3 data, which is about a thousand weather stations, and we might adjust it for that in the future, but right now we're still using 248 because that's the way the scoring tool uses weather data right now; it uses PMY2. I hope that was clear, but it might be clearer later if I show you another slide. So basically we realized that the 19 climate zones were not sensitive enough to the variation in climate within those zones, so that's why we moved -- to this other system. We also found that, you know, if you were in California in a climate that was relatively mild, like in San Diego … I'm trying to figure out how to say this … basically, even though you wouldn't have that much difference, in terms of, you're just going to have less difference among really a crummy house and a really efficient house in California, because you're, or in San Diego I should say, because basically, you just don't have that much heating and cooling needs ‘cause it's very mild. That said, we had to still leave these bins, what we call a bin is like going from a 1 to a 2, or a 2 to a 3, basically the, the energy difference that you can span to be a 2 or a 3 is called a bin in our lingo, at least internally here, and what we found is that if we made the bins too small, like a 12 MB ... or a 10 MBtu or 7 MBtu, we couldn't be accurate enough in terms of predicting exactly where you'd fall, so we had to make the bins wide enough that we could still guarantee or, or be close to guaranteeing that you'd be within one point of accuracy in terms of what you really should be scoring, and we didn't want to make them too big, because if we made them too big it would be very hard to move from, like if you're in, even if you're in a really harsh climate, like in Maine, or in Minnesota, where you could have a huge range, in terms of the energy use from a very efficient home to a very inefficient home, if you made that range too wide, it would be very hard to move from a 2 to a 3 to a 4, because it would just take too much energy change to, to get there; so we had to kind of try to optimize between these different variables that we were trying to address.

Slide 16:

OK, so what we did was, in order to create the, the bins, and I'm just going to move to the next slide because this might make it clearer; I'm not sure it does, but hopefully it does; I'm showing you what we have for New York state, so in New York state it shows that we have, I guess seven different TMY2 weather stations; so one for Albany, Binghamton, Buffalo, whatever, and it could be that there are, you know, places in Connecticut or neighboring states that also use these weather stations, because they're not exactly within a state. It just depends on where you're getting your local weather calculations. And so you can see that, you know, there's not that much difference in some cases, but in New York City it's obviously milder than it is in, I hope I'm saying this right – Massena, which I'm assuming is in, somewhere in the northern part of the state, in upstate New York. So, even though there's not that much variation, there is some variation within, among these seven different weather stations, but in some cases you'll see the same ones; I think four of these actually ended up with the same numbers, but that's not that common, and I just happen to pick New York because I was giving a presentation to someone in New York. And depending on where you are, obviously in California I think there is 19 climate weather stations, and a smaller state, you know Rhode Island has one, there's, there's not a lot. And so what, what we did was we basically ran different homes through the scoring tool; we ran like a really inefficient home, a pretty inefficient home, a moderate home, a really efficient home, through every single one of these, of the zip code associated with each of these weather stations; and that is how we established the range, but when we ran the kind of moderate home, we made sure that, regardless of which weather station you were in within a state, so regardless of where you have a certain house in Albany, Binghamton, Buffalo, wherever, in New York, if you take that same house that we chose as a fairly moderate state, it would always score the same. So, I don't remember if it always scored a 6, or 7, or what it was, but, that's basically how we kind of calibrated our scoring system to the different weather in the country. If you have any questions on that, let me know. OK.

Slide 17:

OK, so now getting back to, kind of what the analysis was that we did, we did two different kinds of, types of Monte Carlo analyses, or, not types, but two different phases of Monte Carlo analyses where we looked at what the effect of asset uncertainty was on the score, and then we also looked at what the effect of behavior was on predicted energy use, or on actual energy use I should say I guess, and what we found was that if you varied, I think as I said earlier, if you varied the key energy assets like insulation, or heating efficiency, etc., what we found was that you could still 90% of the time score a home within 1%, I mean within one point, I'm sorry, of its actual score, so let's say you decided a home, you took 1930s homes, 1980 home, and a 2005 home, and we, then we did the Monte Carlo analysis, which basically created a probability distribution for each one of those homes, we modeled, we did 500 runs for the 1930, the 1980, and the 2005 home where this probability distribution allowed us to perturb these different asset characteristics of the home. And with that, then we found, you know, what

was the range of energy predicted and then how did that affect the score in each of these different climates, so we chose six different places in the country; didn't do it in every single weather station. And 90% of the time we found that, the outcome, in terms of the score was within one point of what we'd actually assigned as the actual value for that home. So let's say we said a home had R-11 wall insulation, and we varied it from R-5 to R-19, or something like that; if we went with the actual, what we had given as the nominal value of this asset, let's say it was supposed to be a six, then 90% of the time it was predicted to be a 6, a 5, or a 7. OK.

{Webinar Host 2: So Joan, we're going to take a break; we have about five or six questions that we can take.}

OK.

{Webinar Host 1: OK, so would it be fair to say that your analysis is somewhat behavior agnostic from the consumer perspective? In other words, is it more building specific?}

If I'm understanding the question correctly, I think my answer is yes. We're not trying to predict a person's behavior, and therefore impact on the energy use of that home, in their home, we're trying to predict how efficient a home is, on average; so we hold the, we are agnostic about behavior, as of that person said.

{Webinar Host 1: OK. Does it try to estimate actual energy consumption which is impacted by energy saving behaviors, which varies across [unknown] and residents?}

No, it does not try to do that.

{Webinar Host 1: OK. In regards to the weather data for each location, is it taken over a period of how many years on average, and is it a daily average, a monthly average, or how is it figured out?}

It's, TMY2 looks at the 30-year history of weather, in that, specific, from that specific weather station, and the reason we would move to TMY2, TMY3, I'm sorry, is that it's more, it's a more recent 30 years, I

forgot what the 30-year span is, for TMY2, but a lot of software tools are moving to TMY3, or some have already done so I believe, and we're hoping to do that in the next iteration of the tool. I believe that then, it does assign it based on, it looks at how many heating degree days, and cooling degree days and all that, to do the modeling. Dave, do you want to say anything about this - in terms of how weather is dealt with?

{Webinar Host 1: OK}

Maybe not. I don't know if Dave's muted or not, but …

{Webinar Host 1: OK, so let's go and move on to another question. Besides correlation with climate zone differences, what other factors were leading factors for differences in energy use? Would it be age, size, construction type, etc.?}

That's interesting, I think we still need to do more analysis there, but what we found when we did the modeled Monte Carlo effort, where we looked at the 1930s, 1980s, and 2005 home, it, it really, I mean definitely age had an impact because we were assigning a more realistic level of, you know , insulation in those homes but, since you know, we weren't saying that people still had boilers from the 1930s, we did obviously make some predictions about, you know, things being updated; there wasn't as much variability in terms of vintage as you might expect; size does make a difference, again it doesn't have to make a huge difference, but it, if you're talking about a huge difference in size, yeah it will make a difference, so if you're comparing a, you know, 5000-square-foot home to a 1500-square-foot home, there will definitely be a difference, because you just have to heat or cool a much larger area.

{If the whole purpose of the tool is to score homes on a relative sense, not based on the calculated energy use, why do you need to run simulations? And then it says you could just generate a score based on the home profile attributes.}

I hope Dave can actually unmute himself because, I don't know if I understand the question, I mean I can start before Dave can chime in, but we did simulations because otherwise there's no way to know how we would create a range for homes, you know, if I, if I'm using the scoring tool and I predict that a really inefficient home is gonna use 350 million BTUs in Minneapolis, but a really inefficient home in Atlanta is gonna use 280 million BTUs, I have to create a range based on that climate, I mean the different climate differences, and then I also have to show that, you know a more, a slightly more efficient home,

regardless of where it is, is gonna score, you know a 2 or a 3, and on down the road. So that's how we did simulations, because we had to look at how changing, you know, R value from R-11 to R-39 in the same home would change the predicted energy use, and therefore correspond to a different score. OK, go ahead, sorry?

{Webinar Host 1: OK. This says, variation in inputs causes less than a 1 point spread, does the tool provide decimal scores, for instance 5.5 versus, or excuse me, 7.5 versus 7.6, or is it only 7 or 8?}

It's only 7 or 8, and that's because given the amount of data that we were collecting for the home, and the lack of diagnostics that we're requiring, because we're really trying to go for a simplified type of assessment, we didn't feel like we could get to that level of granularity, so that's why we don't do that.

{Webinar Host 1: OK. How are variations in building codes and standards taken into account in the analysis?}

Well, we, for instance in the Monte Carlo analysis where we were assigning the different asset values, so you know, we came up with what we thought were reasonable assumptions about a house that was built in 2005 versus '85, versus 1930, and we started with that as a starting point, and then we changed things, saying that, you know, either it could have been built at, you know, better than code, or perhaps it had, you know, maybe we'd gotten the year wrong, and therefore it was a little bit worse than code, so we started with code as a, a basis for making certain assumptions about the type of equipment that would be in the house of that vintage.

Slide 18:

OK, so, in terms, we also, in addition to the Monte Carlo analysis that we did on asset uncertainty, we did some perturbations based on behavior, and for this we actually didn't use the Home Energy Scoring Tool because we did it before the scoring tool was redone, we, and we knew we were going to be changing it during the time of the, of the pilots, so we used the BEOP tool, which is a model that was created by NREL and, and we didn't think it really mattered which tool you were using at that point because you were just, we were just trying to assess how much does behavior really account for energy use variability and what we found that, was that it was pretty significant even if we made fairly conservative assumptions about how much behavior would change and, and we have that documented as to, like, like exactly the same way we did the perturbations for asset, we, in this case we kept the asset exactly the, you know, constant, and then we changed, we only looked at two housing types; we

didn't look at the 1930s, the 1980s, the 2005 homes, we looked at two different houses and we perturbed, or changed the different variables that contribute to, that are associated with occupant behavior, so, how you set your thermostat, how many loads of laundry you do, that sort of thing. How many, you know, how much plug load you're using, etcetera, and what we found was that, again, it was similar in the sense that, well actually 80% of the time on average you would be within 1 point of variance, so, and since we're not trying to do anything with behavior, I think what we're trying to say here is that, if you get a 5, and you're a really, you know, careful user of energy you might be performing more like a 6, in terms of how much energy you're using; but if you're not very good about watching how you use energy you might be more like a 4, so in terms of actual energy use; but your house as an asset would still be a 5. I hope that's sort of clear; and behavior did make a bigger difference in terms of having an effect on energy use in temperate climates like L.A.

Slide 19:

OK, I've already said this a million times, that we're not trying to predict utility bills, but here are some statistics associated with, when we did some runs, we had about 500 REMRate files, actually we had more than that but these are the ones where we had data on both natural gas and on electricity, and so for these 537 homes we ran the home characteristics through the Home Energy Scoring Tool, through SIMPLE, and through REMRate. We also had utility bills for these homes, and NREL basically did analysis where we looked at, you know, how do they compare the actual utility bill with what was being predicted by each of these tools. And you'll see that they were pretty comparable in a lot of ways, and so I think this just gave us sort of a sense of, you know, comfort with the fact that we're predicting pretty much as well as anyone else out, you know, these two software tools in this case, and again, I don't any of these are intending to be predicting energy use, so that's not really the point here; the point is that we're in, you know, kind of, within a reasonable range of actual use; so what you'll see here is that 20, so 61% of the time the Home Energy Scoring Tool's predicted energy use for a home was within 25% of its utility bill use, as you know, is stated on it's utility bill. And so in that case we were, you know, the best, in this case; in the case of within 50% of actual energy use, so predicting, you know, within 50% of what was actually used in the home, SIMPLE was better and we were in the middle.

Slide 20:

OK. This is just a plot, and actually I had some trouble with my, with my abilities today, so it doesn't look, it might be a little bit weird looking here, but, but you'll see that basically if you ran a line, which I couldn't figure how to do, from here, bas.., let's see, if you ran a line so that the 50 was, a 50 touched the 50 and 100 touched the 100, you're, in terms of x y variables, that would be a perfect line of agreement, and what you'll see is that in general, you know it looks like SIMPLE was underestimating relative to actual energy use, and REMRate was overestimating relative to energy use and, basically they all performed sub-lowly, is what I would say; but you could talk to statisticians about what it all means, but in general what I took from it was that we were doing fine. And what I also don't have on here is

that when we ran this information with just Home Energy Scoring Tool before we made the improvements we were higher, like we were overestimating, so I think because of the work that NREL did, and LBNL did, we were able to improve the scoring tool, and so it's, it's more in the, kind of, realm of reality than it was before; before it was doing a little bit of overestimating. OK, on average I should say, because this is all on average, we're not taking into account people's behavior here; we don't know how many people are living in each of these houses, and how they were, you know, how many TV's they were leaving on all the time.

Slide 21:

OK, so in terms of the affordability, this is one of our guiding principles. We found that, as I said, it can be done in a reasonable amount of time. We also confirmed that we don't need to use a blower door test, because it doesn't really affect the score, in terms of, given the lower level of granularity that our score has, in terms of being 10 points rather than 100, or however many points you want to say, we felt that it didn't have that big an impact. And I can show you on the next slide; and I should point out that the scoring tool's free to qualifying partners and assessors, so there's no cost to participate in the program, so all that is, really the cost is the time of the person going into the home.

Slide 22:

Or that's the main cost I should say, there's also cost associated with quality assurance. So, in terms of the blower door test, basically out of the thousand homes that were scored by the pilot partners, a little more than a thousand, 650 about that, had blower door tests done on them, and so we had that data, and so when we reran all of the data from those homes using the blower door test versus not using it, what we, what we found was it really didn't affect the point change terribly much, in terms of where you landed on the scale. If, basically if you don't put in a CFM number from a blower door test the scoring tool will generate a CFM number based on whether or not you say it's sealed or not sealed, in your qualitative opinion, and then also looking at vintage, and other characteristics of what you've entered into the, into the tool; I think it takes into account four or five characteristics and comes up with a default CFM number, because of that. And so what we found was that if you always chose, whatever the qualitative option was, sealed or not sealed that was more in line with what the actual CFM number was from the blower door test, it created on average a quarter point change; and if you always chose the less likely, or the worst option, which means that it was really pretty well sealed but you chose that it was unsealed, or, it wasn't sealed and you chose it was sealed, then on average it would be point 7-point change. So, needless to say, when you did all these different runs, 59% of the time there was no change in the score, and 34% of the time there was 1-point change; so, whatever the difference is there from a hundred, there was more than 1 point, but we didn't feel that it was significant enough to require a blower door test. That said, you can put the blower door test number in if you are doing one anyway.

Slide 23, 24:

OK, in terms of next steps, I'll go through this very quickly so that we have time for more questions, we're implementing the program, recruiting additional partners, we will be updating the scoring tool once a year and releasing it hopefully in January of each year, and we are evaluating assessor requirements; so who can do the assessments, and how we should deal with certification or testing and all that; so we're, we're re-evaluating how that should be done, but for now we're relying on the certifications we've mentioned previously, and on the testing we're using today,…

Slide 25:

… that is for assessors. In terms of future analysis we, you know, continue to want to improve what we're offering, and so we're gonna look at, as some people may have noted in some of their questions, are there certain kinds of, types of homes where we're gonna have more variability and therefore a likelihood of a greater than 1-point variation, are there different characteristics that we should be adding to the scoring tool like PV photovoltaics or swimming pools, or something else that we've left out, and how do we, how did we do in terms of setting the 10-point scale for each of the 248 weather stations, and so that I think we'll only know after we score many, many more homes in the different locations throughout the country; and we're obviously not going to be in all 248 weather station locations immediately. A really important piece of this that we did not get to test in the pilot phase is how does the score change, or does it change how people behave in terms of investing in energy efficient sensible improvements to their home, and if they're doing so, or even if they're not doing so, what can we do to improve that impact, and make it even a greater effect, in terms of getting people to *Slide 2:*
To date based on the pilot findings, tell you a little bit about the analysis that we did over the last year or so, partly from the pilots and then also through another piece of analysis that NREL did for us using model data and also utility bill data. And then, finally I will tell you a little bit about our next steps and what were are planning as we move forward with implementation, both in terms of implementation, but with an emphasis of what we plan to look at further as we proceed.

Slide 3-4:
OK, so some of you may have already heard of these or seen these many times, these are the guiding principles DOE came up with when we put together the Home Energy Score program and we look back at these as kind of our guiding posts to make sure we are still in line with what it is we were trying to achieve initially when we created the program, and I am happy to say that I think for the most part we have been able to fulfill these principles and adhere to them pretty well. The last one about quality assurance, we haven't tested that as closely and it is something that we are going to look at further in this first phase of implementation. We just didn't have the time and range of scoring that we were able to do, in terms of that in the first phase, during the pilot phase. That is something we will look at further.

Slide 5:
So again, giving you a really quick background. The basis for the Home Energy Score is done by collecting about 40 data points in the home and looking at the envelope and the major energy systems, being heating, cooling and hot water. We make some assumptions about how the home is being operated in terms of people that live there, their behavior, and how they set their thermostat, etc. The size of the home does matter, it is used to, the volume is basically used to calculate what the load requirements will be for that home, in addition to obviously looking at where it is located in the country. We do not adjust the 10-point scale for size. The 10-point scale is, however, adjusted for the local climate and we went from 19 climate zones that we used to use for the pilot phase to adjusting that scale for 248 different weather stations around the country, at this point.

Slide 6:
Before we announced the Home Energy Score we did a fair amount of analysis and focus groups, etc., to create the Home Energy Score and the Scoring Tool and then we've done analysis since that time during the pilots and beyond. So, I am going to go through quickly what we found in all that analysis and stress the major findings.

Slide 7:
Before the pilots, that is in 2010, we did hundreds, I guess thousands I should say, of simulations where we looked at different variables, in terms of asset variables that can affect how much energy we are using in the home. So, not looking at how you set your thermostat, or how many loads of laundry you do, but looking, or how many TVs you have, but looking at the assets—the envelope, the windows, the roof, the foundation, insulation levels, the ceiling, whether the ducts are sealed. Things like that, in addition to heating, cooling and hot water, and we isolated those factors that are asset factors and we believe, those we found contribute 2% or more to overall energy use. I should say there are certain things we excluded, like pools, there were just various reasons we decided to do that but that is something we might rethink at a future date. But obviously, pools and hot tubs do contribute, often more than 2%, but it also depends a lot on the behavior of the occupants and how they are using those assets, so that is why we, at that point, left those out, or have to date.

We also found through our focus groups and also through a lot of social science research that had already been done by others, that people are often or frequently motivated by their peers and not necessarily only motivated by dollars, or comfort, or other things they might think about as motivational, that they actually often affected in a kind of illusory way by just what is going on around them. We also found in our focus groups that people responded much better to information that was clearly presented, which seems obvious, but also that was just not an overload of information, that they tended to get distracted if you have them too much to look at upfront. That's not to say you can't give them additional information at the follow-up, but in terms of grabbing people's attention it made a lot of sense to give them simple graphics and information that they could quickly assimilate.

Slide 8:
In terms of what we studied since starting the pilots in 2011, we were looking again at these guiding principles, at what we could try to prove out, that we are actually achieving these, and

seeing how effectively we are achieving these, and we found that when we rescored homes by different assessors, this is not a huge sample but when rescored homes by different assessors, generally they came within 10 percent of each other and only in one case were they more than 1 point off from each other. That is, like if the first person got a 6, the second person got a 5, or a 6, or a 7. So, in general, we thought all of the assessors we were using during the pilots were consistent with each other, in terms of the rescoring number, but that's something we plan to look at further in the next phase. We also found through a modeling effort that, we felt that give the type of scale we had selected, where we don't have a very granular scale, it's only 10 points, we found that if you vary the different factors that are considered asset factors in your modeling that the estimates, even if you are not certain about certain characteristics like wall insulation or attic insulation or how tight the house is, that as long as you were within a reasonable understanding of what's going on in that house, that you could, that you would score the home within one point of what it actually should be scored and I'll go into what I mean by that a little later.

We also did some analysis, NREL did some analysis where we looked at that the scoring tool predicted, in terms of energy use given certain home characteristics versus what was actually going on in that home, where we did have utility bills and I'll show you how we compared to other tools in estimating energy use. But I should stress that this tool is not intended to estimate energy use, and neither were the other tools we looked at. These were asset tools. With that said, it is sometimes comforting to know you're going to be within a reason range of what is really going on there in reality, that is what people are actually using.

Slide 9:
We found that the Score was easy to understand. We did some additional analysis with homeowners and it was easy to understand. We also found that it can be done in what we think is a reasonable amount of time. Fifteen minutes if you are doing it a part of another effort. An hour or less if you are doing it as a stand-alone effort and we also found through sensitivity analysis that we did not believe we needed to do a blower door test. Again, I'll go into more detail on these in other slides.

Slide 10-11:
OK, so in terms of the program improvements to date, as I said yesterday and you probably heard in the past, we had these pilots around the country.

Slide 12:
They led to a lot of improvements as well as other analysis that we were doing with our experts that led us to make adjustments to the tool and other pieces of the program. So, what they told us was that we needed to provide some additional materials for assessors and for homeowners, so we're in the process of developing those or we've filled some of them. We also found that it would be helpful to even go further in simplifying the graphic that we were using during the pilot, so we did make that a bit simpler. We did some analysis that showed that people tend respond to bigger numbers, so we changed the front page of the score from showing just one year of savings to 10 years of savings. The reason we went with 10 years is we thought that was a reasonable amount of time, plus if people wanted to see what they were going to save in a year it was easy to do the math. We are trying to move forward with some type of customization of the information. That is our intent. It has been a little bit tricky just in terms of technically doing it

from the computer side, but we're working towards making that more feasible in the future. We also found that in some cases of the recommendations we were generating during the pilots were off, because they were looking at, for instance they were looking at if you had, you know, a 20-year-old boiler and you were gonna go replace it, we were calculating the savings as, what you would replace it if you bought an ENERGY STAR boiler or furnace, or whatever, today, versus a off-the-shelf non-ENERGY STAR. It wasn't looking back at the base case of the house. So this was messing things up, both in terms of underestimating savings, if you had a crummy piece of equipment, because it was far worse than what you could get, even if you didn't buy ENERGY STAR, or it was overestimating savings because we weren't taking into account that you already had an ENERGY STAR heater or cooling system. So we've fixed that, and we also found that, in some cases there were some pilot partners that really stressed the fact that they weren't, they were a bit perturbed by the fact that if you had some really poor scoring homes, that they could still not necessarily get to a be a 7, 8, or 9, or 10, that they might only get to be a 6, or in some cases possibly even just like a 5, or a 4, and so they felt that it was important to provide material that would allow people to understand the context of the scale, and that it's, well, it allows you to compare homes, it's also useful as a point to understand that if you're a 2 today and you can get to be a 6, that's still, you know, great, in terms of what's cost effective for that home. We don't have enough data at this point to show you that if you're a 1940s bungalow, you know, the best you're gonna be is a 7, or, I making that up, but that is the kind of information that I think people would find useful; we don't have it yet, ‘cause we just haven't scored enough homes. We did a lot of work with the assessors, the 31 assessors who were involved in the pilots, they gave us tremendous amount of feedback on, kind of what they thought of the scoring tool, and the program in general. We added a lot of information within the scoring tool so you could click on question marks for every single field, and get additional guidance that way to guide the assessor as they're going through it in case they forgot what we're actually asking for in a specific field. They also indicated where we might not have enough choices, or too many choices, and so we, we did make some changes to the scoring tool based on that. We've greatly enhanced our training and our testing to cover not only the scoring tool, but also to, and the scoring program, but also to cover some building science issues as well, just to be as a refresher and to ensure that people have been going into homes, you know relatively recently, and know how to make some basic calculations. We've also expanded the testing to include, not just the multiple choice portion which it was before, we expanded that but we've also included a second part of the test which requires the assessor to make some calculations about homes and then enter that and score some fictitious homes in the scoring tool itself.

Slide 13:
We looked at three different software tools. We looked our own Home Energy Scoring tool, we looked at REMRate, and we looked at SIMPLE, a version of SIMPLE, and all of those were intended to provide an asset, well, I mean they can be just used for other things, I don't want to speak for the other tools, but, in, none of them were intended that the way we were using them to predict energy use in terms of actual utility bills, but we did that kind of analysis where we compared those estimates to actual energy use, just to give us a sense of, you know, grounding, of whether or not these things were way off, or whether they were in a reasonable range.

Slide 14:
OK, so, getting back to, OK, so in terms of the scoring tool improvements, I already mentioned

some of the ways that we changed how we were calculating savings and generating recommendations, I think we did some improvements to the user interface, I also mentioned some of that. I not gonna go through every single one of these but we will make these slides available if you're interested in, in seeing the specific details, again there were certain things that assessors indicated as being just either inaccurate or possibly not including enough options for them, and so we made those adjustments. We also adjusted the, the plug load -- can now be scaled by floor area, it used to be just a sort of standard number for every house, and now the lighting and small appliances for the home are adjusted in terms of the load calculations to be scaled by floor area. We did change the standard occupancy assumption based on some research that Danny Parker did. We found that, in general there was a sort of a, a linear distribution, a linear relationship between the number of bedrooms and actually how many people are living in a house, it wasn't a direct one-to-one, but I can make that equation available to people if they want, I think it was like .7 plus .3 times bedrooms or something, and so that's what we used. It didn't change too much, in terms of what ended up being the case of how many people we're assuming are living in a house, but the number of occupants is tagged to the number of bedrooms… or tied to. We made some adjustments to our defaults, and some of that was done based on information that we got out of the RECS 2009 data, as well as information that was updated from the Building America house simulation protocols and benchmark definitions. And I also mentioned that we had gone from 19 REC zones that we were using for climate to 248, and that was partly by, because of something that the folks at CEC in California raised to us and it really it turned out to be important to make that change for the whole country, and I can go through that analysis as well.

Slide 15:
{Webinar Host 1: Will you take a couple questions?}

Sure.

{Webinar Host 1: OK, can you address the decision to go from 1 to 10 with 10 being the best score, opposite of the HERS rating, and the chance that this could create confusion with the various rankings?}

I think that's something that people have struggled with, as to you know, whether we're trying to push net zero, or we're trying to show zero energy use; the reason we went with the 10-point scale is that we though in general that people outside of the energy world understand a 10-point scale, like, you know, going back to grade school and taking 10-point quizzes, or seeing Bo Derek in the movie "10." Ten was associated with good, or great, and 1 was associated with not-so-great. And that's why we went to that. I don't think there's any, you know, perfect solution to what scale you should use, but that's how we came to that.

{Webinar Host 1: OK, and could you clarify your comments in terms of the other tools not being designed to test against utility bills?}

I think I misstated that, if that's the impression I gave that's not what I meant. I meant that, I was saying that we are not expecting the Home Energy Scoring Tool to predict your utility bills, but we did an analysis where we compared the estimated energy use from the Home Energy Scoring

Tool to actual energy use from homes. We did the same thing with two other tools. It's not that they were, you know I'm just saying, it's not intended to do that, but it was, we were kind of trying to do a comparison by looking at three different tools and how well their predicted energy use corresponded to utility bills. That's all, it's just, I was trying not to overstate that ours was not intended and others were, it's, I think in all cases, and maybe Dave can answer this, in all cases those were being used as asset tools, and we held certain assumptions constant about how they were being used, because we did not have the information about who was living in those homes. And Dave, if you want to add anything you can.

{Webinar Host 1: OK, well then, could you clarify what is the accuracy level compared to a real energy bill?}

That's a good question. I think accuracy is all within the eye of the beholder, I guess. What we're trying to achieve is accuracy in terms of relative accuracy. We're not saying that your home is going to use this much energy regardless of who lives in it. Obviously, if you're a person who leaves the lights on and, you know, does three loads of laundry every day, and you know, washed the dishes constantly, you're gonna have a higher energy bill than somebody who's more careful, and that's just the way it is. So we're not trying to predict how much energy is going to be used by a specific homeowner, so we're not accurate in that sense. What we are accurate, and I think we can feel very confident about this is that if you're, if you take the same homeowner and you put them in two different homes, if one home has a not-so-great heating system and cooling system and it's not very well insulated, that homeowner is going to use more energy in that home than in the one that scores, if that home scores a 3, and another home scores a 7, if they live in the 7 and they live the same way as if they were going to live in the 3, they will use less energy. So it gives you a point of comparison between homes, but it does not predict the actual utility bills, so there's no accuracy involved in that sense. It's accurate in terms of scoring homes in a relative sense. A crummy home will get a poor score, a medium home will get a medium score, and a really good home will get a good score.

{Webinar Host 1: OK, thanks, Joan.}

OK, so in terms of going from 19 climate zones to the 248, the reason we did that is that we, we took, I have to remember how we did this, we basically took the same home and used different zip codes within, like let's say we were looking at Zone 1 from the 19 climate zones we used to have from RECS, I believe that was like the NorthEast; so if we took two different zip codes from Maine and Massachusetts, if they were all on the same climate zone, and we modeled the same home in those two different zip codes, what we found was they wouldn't necessarily score the same, and obviously that was a problem, because it turned out that even if you were within the same REC zone, which was, you know, a pretty big zone in some cases, you'd have different weather obviously if you're in Bangor, Maine, then if you're on the Cape, or if you're in western Mass., or, wherever you are; you could be in the same REC zone but you have different weather, and the load calculations done by the tool are based on the local weather, but previously we'd only created 10-point scales that were adjusted to 19 different areas of the country. Now they are adjusted based on the local weather so, I should say it's only based on the 248 weather stations associated with PMY2 data, now we've moved, or the country's moved to PMY3 data, which is about a thousand weather stations, and we might adjust it for that in the future, but right now

we're still using 248 because that's the way the scoring tool uses weather data right now; it uses PMY2. I hope that was clear, but it might be clearer later if I show you another slide. So basically we realized that the 19 climate zones were not sensitive enough to the variation in climate within those zones, so that's why we moved -- to this other system. We also found that, you know, if you were in California in a climate that was relatively mild, like in San Diego … I'm trying to figure out how to say this … basically, even though you wouldn't have that much difference, in terms of, you're just going to have less difference among really a crummy house and a really efficient house in California, because you're, or in San Diego I should say, because basically, you just don't have that much heating and cooling needs 'cause it's very mild. That said, we had to still leave these bins, what we call a bin is like going from a 1 to a 2, or a 2 to a 3, basically the, the energy difference that you can span to be a 2 or a 3 is called a bin in our lingo, at least internally here, and what we found is that if we made the bins too small, like a 12 MB ... or a 10 MBtu or 7 MBtu, we couldn't be accurate enough in terms of predicting exactly where you'd fall, so we had to make the bins wide enough that we could still guarantee or, or be close to guaranteeing that you'd be within one point of accuracy in terms of what you really should be scoring, and we didn't want to make them too big, because if we made them too big it would be very hard to move from, like if you're in, even if you're in a really harsh climate, like in Maine, or in Minnesota, where you could have a huge range, in terms of the energy use from a very efficient home to a very inefficient home, if you made that range too wide, it would be very hard to move from a 2 to a 3 to a 4, because it would just take too much energy change to, to get there; so we had to kind of try to optimize between these different variables that we were trying to address.

Slide 16:
OK, so what we did was, in order to create the, the bins, and I'm just going to move to the next slide because this might make it clearer; I'm not sure it does, but hopefully it does; I'm showing you what we have for New York state, so in New York state it shows that we have, I guess seven different TMY2 weather stations; so one for Albany, Binghamton, Buffalo, whatever, and it could be that there are, you know, places in Connecticut or neighboring states that also use these weather stations, because they're not exactly within a state. It just depends on where you're getting your local weather calculations. And so you can see that, you know, there's not that much difference in some cases, but in New York City it's obviously milder than it is in, I hope I'm saying this right – Massena, which I'm assuming is in, somewhere in the northern part of the state, in upstate New York. So, even though there's not that much variation, there is some variation within, among these seven different weather stations, but in some cases you'll see the same ones; I think four of these actually ended up with the same numbers, but that's not that common, and I just happen to pick New York because I was giving a presentation to someone in New York. And depending on where you are, obviously in California I think there is 19 climate weather stations, and a smaller state, you know Rhode Island has one, there's, there's not a lot. And so what, what we did was we basically ran different homes through the scoring tool; we ran like a really inefficient home, a pretty inefficient home, a moderate home, a really efficient home, through every single one of these, of the zip code associated with each of these weather stations; and that is how we established the range, but when we ran the kind of moderate home, we made sure that, regardless of which weather station you were in within a state, so regardless of where you have a certain house in Albany, Binghamton, Buffalo, wherever, in New York, if you take that same house that we chose as a fairly moderate state, it would always score the

same. So, I don't remember if it always scored a 6, or 7, or what it was, but, that's basically how we kind of calibrated our scoring system to the different weather in the country. If you have any questions on that, let me know. OK.

Slide 17:
OK, so now getting back to, kind of what the analysis was that we did, we did two different kinds of, types of Monte Carlo analyses, or, not types, but two different phases of Monte Carlo analyses where we looked at what the effect of asset uncertainty was on the score, and then we also looked at what the effect of behavior was on predicted energy use, or on actual energy use I should say I guess, and what we found was that if you varied, I think as I said earlier, if you varied the key energy assets like insulation, or heating efficiency, etc., what we found was that you could still 90% of the time score a home within 1%, I mean within one point, I'm sorry, of its actual score, so let's say you decided a home, you took 1930s homes, 1980 home, and a 2005 home, and we, then we did the Monte Carlo analysis, which basically created a probability distribution for each one of those homes, we modeled, we did 500 runs for the 1930, the 1980, and the 2005 home where this probability distribution allowed us to perturb these different asset characteristics of the home. And with that, then we found, you know, what was the range of energy predicted and then how did that affect the score in each of these different climates, so we chose six different places in the country; didn't do it in every single weather station. And 90% of the time we found that, the outcome, in terms of the score was within one point of what we'd actually assigned as the actual value for that home. So let's say we said a home had R-11 wall insulation, and we varied it from R-5 to R-19, or something like that; if we went with the actual, what we had given as the nominal value of this asset, let's say it was supposed to be a six, then 90% of the time it was predicted to be a 6, a 5, or a 7. OK.

{Webinar Host 2: So Joan, we're going to take a break; we have about five or six questions that we can take.}

OK.

{Webinar Host 1: OK, so would it be fair to say that your analysis is somewhat behavior agnostic from the consumer perspective? In other words, is it more building specific?}

If I'm understanding the question correctly, I think my answer is yes. We're not trying to predict a person's behavior, and therefore impact on the energy use of that home, in their home, we're trying to predict how efficient a home is, on average; so we hold the, we are agnostic about behavior, as of that person said.

{Webinar Host 1: OK. Does it try to estimate actual energy consumption which is impacted by energy saving behaviors, which varies across [unknown] and residents?}

No, it does not try to do that.

{Webinar Host 1: OK. In regards to the weather data for each location, is it taken over a period of how many years on average, and is it a daily average, a monthly average, or how is it figured out?}

It's, TMY2 looks at the 30-year history of weather, in that, specific, from that specific weather station, and the reason we would move to TMY2, TMY3, I'm sorry, is that it's more, it's a more recent 30 years, I forgot what the 30-year span is, for TMY2, but a lot of software tools are moving to TMY3, or some have already done so I believe, and we're hoping to do that in the next iteration of the tool. I believe that then, it does assign it based on, it looks at how many heating degree days, and cooling degree days and all that, to do the modeling. Dave, do you want to say anything about this - in terms of how weather is dealt with?

{Webinar Host 1: OK}

Maybe not. I don't know if Dave's muted or not, but …

{Webinar Host 1: OK, so let's go and move on to another question. Besides correlation with climate zone differences, what other factors were leading factors for differences in energy use? Would it be age, size, construction type, etc.?}

That's interesting, I think we still need to do more analysis there, but what we found when we did the modeled Monte Carlo effort, where we looked at the 1930s, 1980s, and 2005 home, it, it really, I mean definitely age had an impact because we were assigning a more realistic level of, you know , insulation in those homes but, since you know, we weren't saying that people still had boilers from the 1930s, we did obviously make some predictions about, you know, things being updated; there wasn't as much variability in terms of vintage as you might expect; size does make a difference, again it doesn't have to make a huge difference, but it, if you're talking about a huge difference in size, yeah it will make a difference, so if you're comparing a, you know, 5000-square-foot home to a 1500-square-foot home, there will definitely be a difference, because you just have to heat or cool a much larger area.

{If the whole purpose of the tool is to score homes on a relative sense, not based on the calculated energy use, why do you need to run simulations? And then it says you could just generate a score based on the home profile attributes.}

I hope Dave can actually unmute himself because, I don't know if I understand the question, I mean I can start before Dave can chime in, but we did simulations because otherwise there's no way to know how we would create a range for homes, you know, if I, if I'm using the scoring tool and I predict that a really inefficient home is gonna use 350 million BTUs in Minneapolis, but a really inefficient home in Atlanta is gonna use 280 million BTUs, I have to create a range based on that climate, I mean the different climate differences, and then I also have to show that, you know a more, a slightly more efficient home, regardless of where it is, is gonna score, you know a 2 or a 3, and on down the road. So that's how we did simulations, because we had to look at how changing, you know, R value from R-11 to R-39 in the same home would change the predicted energy use, and therefore correspond to a different score. OK, go ahead, sorry?

{Webinar Host 1: OK. This says, variation in inputs causes less than a 1 point spread, does the tool provide decimal scores, for instance 5.5 versus, or excuse me, 7.5 versus 7.6, or is it only 7 or 8?}

It's only 7 or 8, and that's because given the amount of data that we were collecting for the home, and the lack of diagnostics that we're requiring, because we're really trying to go for a simplified type of assessment, we didn't feel like we could get to that level of granularity, so that's why we don't do that.

{Webinar Host 1: OK. How are variations in building codes and standards taken into account in the analysis?}

Well, we, for instance in the Monte Carlo analysis where we were assigning the different asset values, so you know, we came up with what we thought were reasonable assumptions about a house that was built in 2005 versus '85, versus 1930, and we started with that as a starting point, and then we changed things, saying that, you know, either it could have been built at, you know, better than code, or perhaps it had, you know, maybe we'd gotten the year wrong, and therefore it was a little bit worse than code, so we started with code as a, a basis for making certain assumptions about the type of equipment that would be in the house of that vintage.

Slide 18:
OK, so, in terms, we also, in addition to the Monte Carlo analysis that we did on asset uncertainty, we did some perturbations based on behavior, and for this we actually didn't use the Home Energy Scoring Tool because we did it before the scoring tool was redone, we, and we knew we were going to be changing it during the time of the, of the pilots, so we used the BEOP tool, which is a model that was created by NREL and, and we didn't think it really mattered which tool you were using at that point because you were just, we were just trying to assess how much does behavior really account for energy use variability and what we found that, was that it was pretty significant even if we made fairly conservative assumptions about how much behavior would change and, and we have that documented as to, like, like exactly the same way we did the perturbations for asset, we, in this case we kept the asset exactly the, you know, constant, and then we changed, we only looked at two housing types; we didn't look at the 1930s, the 1980s, the 2005 homes, we looked at two different houses and we perturbed, or changed the different variables that contribute to, that are associated with occupant behavior, so, how you set your thermostat, how many loads of laundry you do, that sort of thing. How many, you know, how much plug load you're using, etcetera, and what we found was that, again, it was similar in the sense that, well actually 80% of the time on average you would be within 1 point of variance, so, and since we're not trying to do anything with behavior, I think what we're trying to say here is that, if you get a 5, and you're a really, you know, careful user of energy you might be performing more like a 6, in terms of how much energy you're using; but if you're not very good about watching how you use energy you might be more like a 4, so in terms of actual energy use; but your house as an asset would still be a 5. I hope that's sort of clear; and behavior did make a bigger difference in terms of having an effect on energy use in temperate climates like L.A.

Slide 19:
OK, I've already said this a million times, that we're not trying to predict utility bills, but here are some statistics associated with, when we did some runs, we had about 500 REMRate files, actually we had more than that but these are the ones where we had data on both natural gas and on electricity, and so for these 537 homes we ran the home characteristics through the Home Energy Scoring Tool, through SIMPLE, and through REMRate. We also had utility bills for

these homes, and NREL basically did analysis where we looked at, you know, how do they compare the actual utility bill with what was being predicted by each of these tools. And you'll see that they were pretty comparable in a lot of ways, and so I think this just gave us sort of a sense of, you know, comfort with the fact that we're predicting pretty much as well as anyone else out, you know, these two software tools in this case, and again, I don't any of these are intending to be predicting energy use, so that's not really the point here; the point is that we're in, you know, kind of, within a reasonable range of actual use; so what you'll see here is that 20, so 61% of the time the Home Energy Scoring Tool's predicted energy use for a home was within 25% of its utility bill use, as you know, is stated on it's utility bill. And so in that case we were, you know, the best, in this case; in the case of within 50% of actual energy use, so predicting, you know, within 50% of what was actually used in the home, SIMPLE was better and we were in the middle.

Slide 20:
OK. This is just a plot, and actually I had some trouble with my, with my abilities today, so it doesn't look, it might be a little bit weird looking here, but, but you'll see that basically if you ran a line, which I couldn't figure how to do, from here, bas.., let's see, if you ran a line so that the 50 was, a 50 touched the 50 and 100 touched the 100, you're, in terms of x y variables, that would be a perfect line of agreement, and what you'll see is that in general, you know it looks like SIMPLE was underestimating relative to actual energy use, and REMRate was overestimating relative to energy use and, basically they all performed sub-lowly, is what I would say; but you could talk to statisticians about what it all means, but in general what I took from it was that we were doing fine. And what I also don't have on here is that when we ran this information with just Home Energy Scoring Tool before we made the improvements we were higher, like we were overestimating, so I think because of the work that NREL did, and LBNL did, we were able to improve the scoring tool, and so it's, it's more in the, kind of, realm of reality than it was before; before it was doing a little bit of overestimating. OK, on average I should say, because this is all on average, we're not taking into account people's behavior here; we don't know how many people are living in each of these houses, and how they were, you know, how many TV's they were leaving on all the time.

Slide 21:
OK, so in terms of the affordability, this is one of our guiding principles. We found that, as I said, it can be done in a reasonable amount of time. We also confirmed that we don't need to use a blower door test, because it doesn't really affect the score, in terms of, given the lower level of granularity that our score has, in terms of being 10 points rather than 100, or however many points you want to say, we felt that it didn't have that big an impact. And I can show you on the next slide; and I should point out that the scoring tool's free to qualifying partners and assessors, so there's no cost to participate in the program, so all that is, really the cost is the time of the person going into the home.

Slide 22:
Or that's the main cost I should say, there's also cost associated with quality assurance. So, in terms of the blower door test, basically out of the thousand homes that were scored by the pilot partners, a little more than a thousand, 650 about that, had blower door tests done on them, and so we had that data, and so when we reran all of the data from those homes using the blower

door test versus not using it, what we, what we found was it really didn't affect the point change terribly much, in terms of where you landed on the scale. If, basically if you don't put in a CFM number from a blower door test the scoring tool will generate a CFM number based on whether or not you say it's sealed or not sealed, in your qualitative opinion, and then also looking at vintage, and other characteristics of what you've entered into the, into the tool; I think it takes into account four or five characteristics and comes up with a default CFM number, because of that. And so what we found was that if you always chose, whatever the qualitative option was, sealed or not sealed that was more in line with what the actual CFM number was from the blower door test, it created on average a quarter point change; and if you always chose the less likely, or the worst option, which means that it was really pretty well sealed but you chose that it was unsealed, or, it wasn't sealed and you chose it was sealed, then on average it would be point 7-point change. So, needless to say, when you did all these different runs, 59% of the time there was no change in the score, and 34% of the time there was 1-point change; so, whatever the difference is there from a hundred, there was more than 1 point, but we didn't feel that it was significant enough to require a blower door test. That said, you can put the blower door test number in if you are doing one anyway.

Slide 23, 24:
OK, in terms of next steps, I'll go through this very quickly so that we have time for more questions, we're implementing the program, recruiting additional partners, we will be updating the scoring tool once a year and releasing it hopefully in January of each year, and we are evaluating assessor requirements; so who can do the assessments, and how we should deal with certification or testing and all that; so we're, we're re-evaluating how that should be done, but for now we're relying on the certifications we've mentioned previously, and on the testing we're using today,…

Slide 25:
… that is for assessors. In terms of future analysis we, you know, continue to want to improve what we're offering, and so we're gonna look at, as some people may have noted in some of their questions, are there certain kinds of, types of homes where we're gonna have more variability and therefore a likelihood of a greater than 1-point variation, are there different characteristics that we should be adding to the scoring tool like PV photovoltaics or swimming pools, or something else that we've left out, and how do we, how did we do in terms of setting the 10-point scale for each of the 248 weather stations, and so that I think we'll only know after we score many, many more homes in the different locations throughout the country; and we're obviously not going to be in all 248 weather station locations immediately. A really important piece of this that we did not get to test in the pilot phase is how does the score change, or does it change how people behave in terms of investing in energy efficient sensible improvements to their home, and if they're doing so, or even if they're not doing so, what can we do to improve that impact, and make it even a greater effect, in terms of getting people to improve the efficiency of their homes; either the ones they own, or the ones they're buying. OK, and would there be ways that we can adjust the scale of the materials to increase that motivation.

Slide 26:
I already talked a little bit about assessor qualifications, are there different things that we should be looking at, not just certification levels, but requiring a certain level of experience, or making

the test a stand-alone rather than requiring certification, and we're also looking at how does it make sense in terms of what kind of scale-up can we do; right now we have 20 plus partners; if this is going to be something that can be done for lots and lots of homes across the country, then we need a way that, you know, DOE doesn't have to be so heavily involved with every partner, so are there ways that we can set up quality assurance programs around the country, are there other kinds of providers that can do so, and we also need to test the use of the application program interface which is the way that different software tools can basically link to the scoring tool so that you don't have to re-enter data. So, those are the kind of test analysis we plan to do over the next year and beyond.

improve the efficiency of their homes; either the ones they own, or the ones they're buying. OK, and would there be ways that we can adjust the scale of the materials to increase that motivation.

Slide 26:

I already talked a little bit about assessor qualifications, are there different things that we should be looking at, not just certification levels, but requiring a certain level of experience, or making the test a stand-alone rather than requiring certification, and we're also looking at how does it make sense in terms of what kind of scale-up can we do; right now we have 20 plus partners; if this is going to be something that can be done for lots and lots of homes across the country, then we need a way that, you know, DOE doesn't have to be so heavily involved with every partner, so are there ways that we can set up quality assurance programs around the country, are there other kinds of providers that can do so, and we also need to test the use of the application program interface which is the way that different software tools can basically link to the scoring tool so that you don't have to re-enter data. So, those are the kind of test analysis we plan to do over the next year and beyond.